ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Laboratorio di

Sicurezza Informatica

Fondamenti di crittografia
Crittografia classica

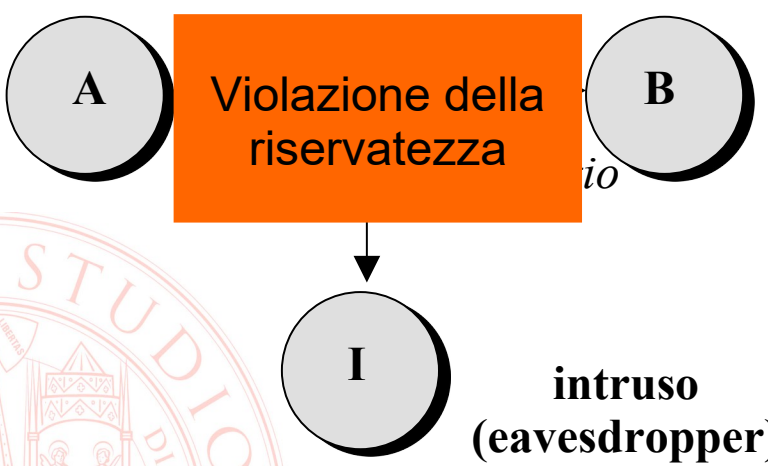
Marco Prandini

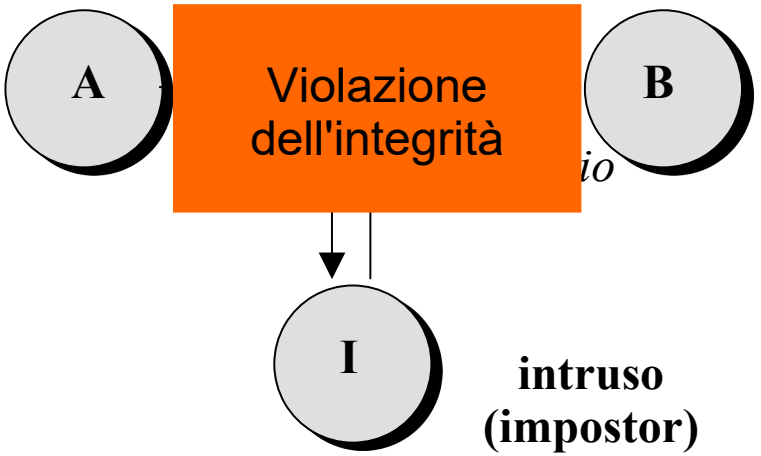
Dipartimento di Informatica – Scienza e Ingegneria

Ricordiamo le sfaccettature della sicurezza delle informazioni

- Confidentiality (riservatezza)
- Integrity (integrità)
  - Authenticity (paternità)
- Availability (disponibilità)

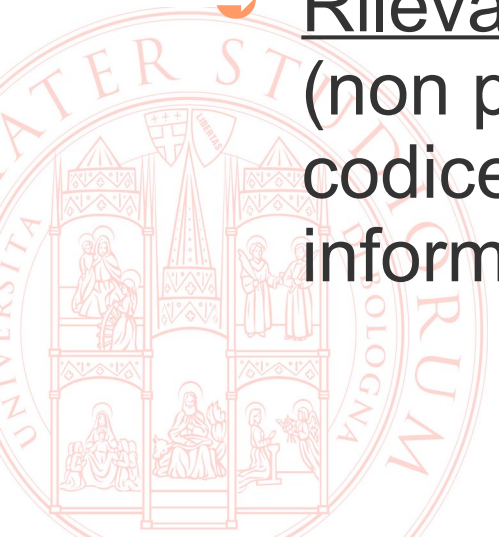
Mondi ideali e reali



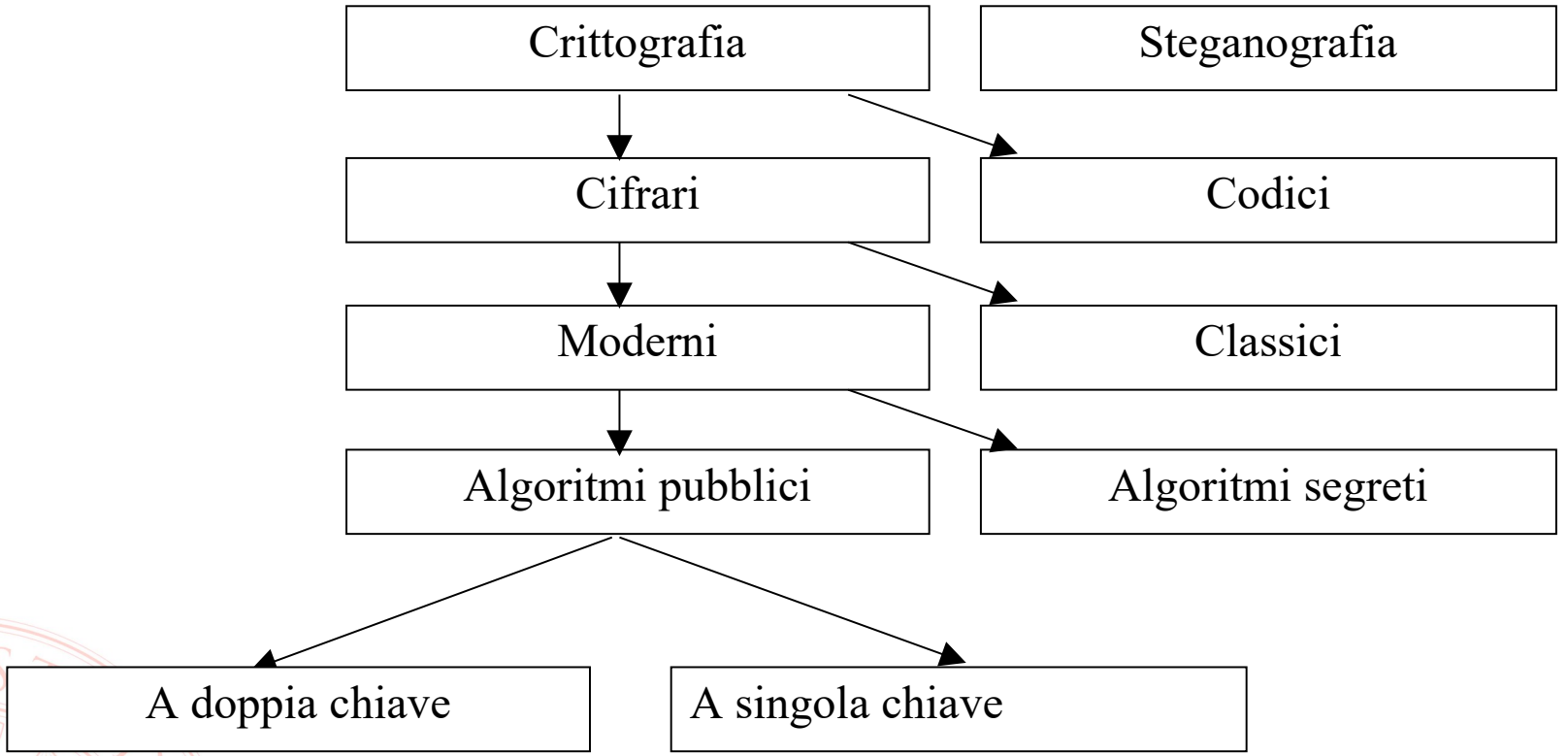


Soluzione: crittografia

- Un'elaborazione matematica e algoritmica della codifica delle informazioni
- Prevenire la violazione della riservatezza (una rilevazione a posteriori sarebbe inefficace!): →alterare il codice in modo da renderlo incomprensibile a chi non ha diritto di apprendere le informazioni
- Rilevare la violazione dell'integrità e autenticità (non può essere prevenuta!) → aggiungere al codice elementi che permettano la verifica delle informazioni ricevute



Tecniche per la realizzazione delle primitive



Una novità introdotta da Internet?

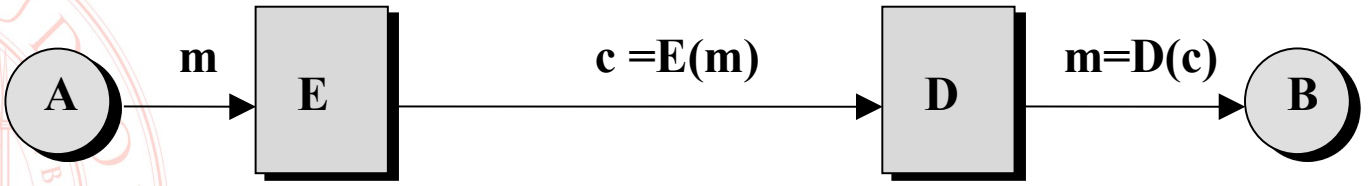
- VI sec. a.C. - Il cifrario Atbash degli Ebrei
  - Sostituzione monoalfabietica
- V sec. a.C. - La tavoletta di Demarato
  - Steganografia
- IV sec. a.C. - La scitala degli Spartani
  - Trasposizione
- IV sec. a.C. - Lo schiavo rapato di Istieo
  - Steganografia
- I sec. a.C. - Il cifrari di Cesare
  - Sostituzione monoalfabetica
- VIII sec. d.C. - Il trattato di Al-Kindi
  - Studio sistematico della crittoanalisi

Steganografia

- L’arte e la scienza del comunicare senza che altri se ne accorgano
- Esempi storici:
  - La tavoletta cerata di Demerato
  - Lo schiavo "rapato" di Istieo
  - Gli inchiostri invisibili
- Tecniche moderne
  - Modifica dei bit meno significativi di dati multimediali

Cifrari per la riservatezza

- Due operazioni
  - Cifratura
    converte il testo in chiaro in testo cifrato
  - Decifrazione
    converte il testo cifrato in testo in chiaro



Codici

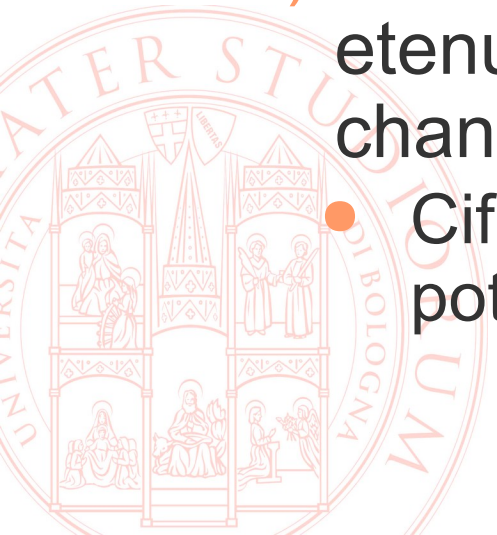
- Sostituzione di stringhe
  - Tipicamente parole
- Limitato dal dizionario
  - Capacità espressive ridotte
  - Complessità di rappresentazione del codice

Algoritmi segreti

- Rappresentano una possibilità di rendere irrealizzabile D
- Benefici apparenti: difficoltà di studiare come invertire la cifratura
- Problemi:
  - mancanza di revisione della qualità
  - difficoltà di diffusione delle procedure
  - difficoltà di sostituzione delle procedure

I principi di Kerckhoffs (1883)

1) Le système doit être matériellement, sinon mathématiquement, indéchiffrable ;
   - Sicurezza *computazionale* o *assoluta*
2) Il faut qu'il n'exige pas le secret, et qu'il puisse sans inconvénient tomber entre les mains de l'ennemi ;
   - ~~segreto=algoritmo~~ segreto=chiave!
3) La clef doit pouvoir en être communiquée et etenue sans le secours de notes écrites, et être changée ou modifiée au gré des correspondants;
   - Cifratura = ricordare un segreto semplice per poter scambiare molti segreti arbitrari



Metodi della crittoanalisi

- A seconda del materiale a disposizione del crittanalista si possono avere diverse opportunità di attacco

- **Forza bruta**
  - Si tira ad indovinare D (o il suo particolare segreto) e si decifra il testo intercettato: se non ha alcun senso si ripete il procedimento

- **Solo testo cifrato**
  - Si eseguono analisi statistiche su una grande quantità di materiale cifrato e se ne usano le indicazioni per individuare quale p probabilmente corrisponde ad un dato c

- **Testo in chiaro noto**
  - Ci si procura in qualche modo sia dei testi cifrati, sia i corrispondenti testi in chiaro e si cerca di dedurre D analizzando le varie coppie

- **Testo scelto**
  - Si può scegliere testo da cifrare o da far decifrare per ottimizzare il procedimento di deduzione della chiave

- **Rubber hose**
  - Si minaccia, ricatta o tortura qualcuno finché non cede la chiave.

In sintesi: crittoanalisi e crittografia

- Di fronte a un testo cifrato con algoritmo noto, cosa può sempre fare un crittoanalista?
  - Analizzare le proprietà statistiche del testo

    → robustezza = capacità dell'algoritmo di **occultare le proprietà del testo in chiaro**
  - Cercare la chiave tra tutte quelle possibili

    → sicurezza assoluta = rendere totalmente **indistinguibile la chiave giusta** dalle altre

    → sicurezza computazionale = rendere **troppo oneroso il processo di ricerca** della chiave

“Mattoni” della robustezza: confusione

- La proprietà di *confusione* misura il grado in cui la struttura della chiave viene resa irriconoscibile nel testo cifrato
  - Una modifica di un singolo elemento della chiave dovrebbe riflettersi sul 50% del testo cifrato
  - L’analisi del testo cifrato non restituisce indicazioni utili sul valore della chiave

“Mattoni” della robustezza: diffusione

- La proprietà di *diffusione* misura il grado in cui le proprietà statistiche degli elementi del testo in chiaro vengono sparse sugli elementi del testo cifrato
  - Una modifica di un singolo elemento del testo in chiaro dovrebbe alterare il 50% del testo cifrato
  - L’analisi del testo cifrato non restituisce indicazioni utili sul testo in chiaro

Sostituzione monoalfabetica

- La sostituzione è il modo più semplice di introdurre confusione
- Cifrario di Cesare, Agony Columns del Times, parole crociate crittografate della Settimana Enigmistica, …

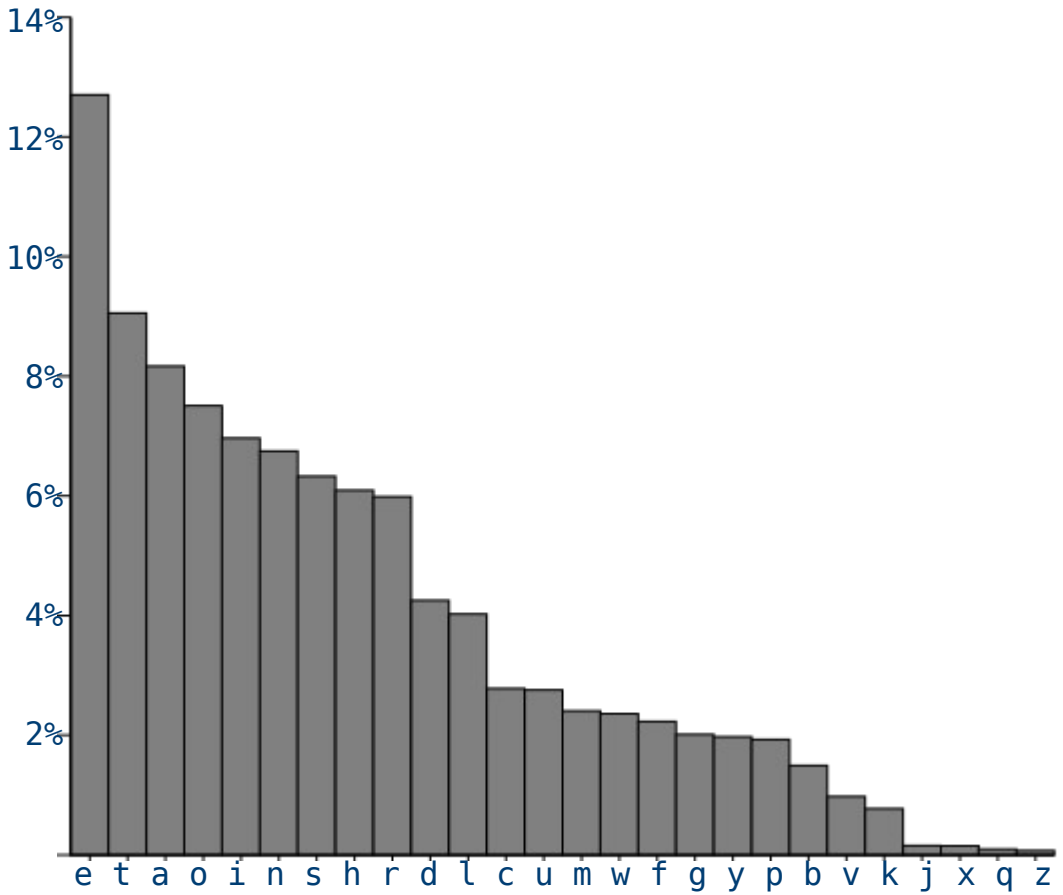
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Q	W	E	R	T	Y	U	I	O	P	L	K	J	H	G	F	D	S	A	Z	X	C	V	B	N	M

- **CRITTOGRAFIA →ESOZZGUSQYOQ**
- Ricerca della chiave: spazio di 26! ≈ $4*10^{26} \approx 2^{88}$
- Robustezza…

Attacco alla sostituzione

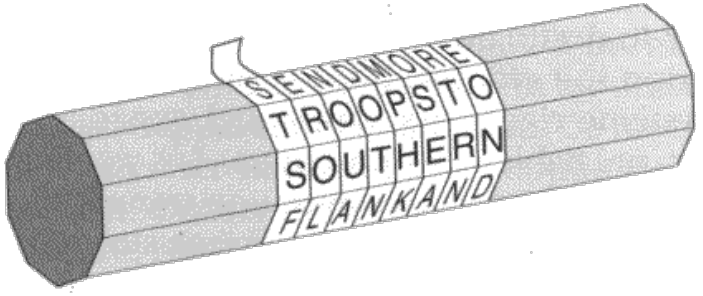
Nel linguaggio naturale, estremamente facile con le statistiche di frequenza dei caratteri (in figura il grafico per la lingua inglese)

Nel mondo binario, la "lettera" può essere un lungo blocco di bit
→ frequenze basse e uniformi (compressione)
→ buona efficacia!



Trasposizione

- La trasposizione è il modo più semplice di introdurre diffusione
- La scitala degli Spartani



- Algoritmicamente basta una tabella scritta per colonne e letta per righe

ALLE PROSSIME ELEZIONI MI PRESENTO ANCHE IO

A	P	I	L	N		E	A	
L	R	M	E	I	P	N	N	I
L	O	E	Z		R	T	C	O
E	S		I	M	E	O	H	
	S	E	O	I	S		E	

APILN EA LRMEIPNNI LOEZ RTCOES IMEOH SEOIS E

Trasposizione

- Ricerca della chiave:
  - dimensione della tabella
  - ordine di lettura delle righe
- Robustezza
  - Statistiche dei *digrammi* e *trigrammi*
    - Permettono di dedurre la dimensione della tabella
  - Per nulla banale se applicata ripetutamente

- Lingua inglese

  TH 3,16%,

  IN 1,54%

  ER 1,33%

  RE 1,3%

  ecc.

  THE 4,72

  ING 1,42

  ecc.

Sostituzione polialfabetica

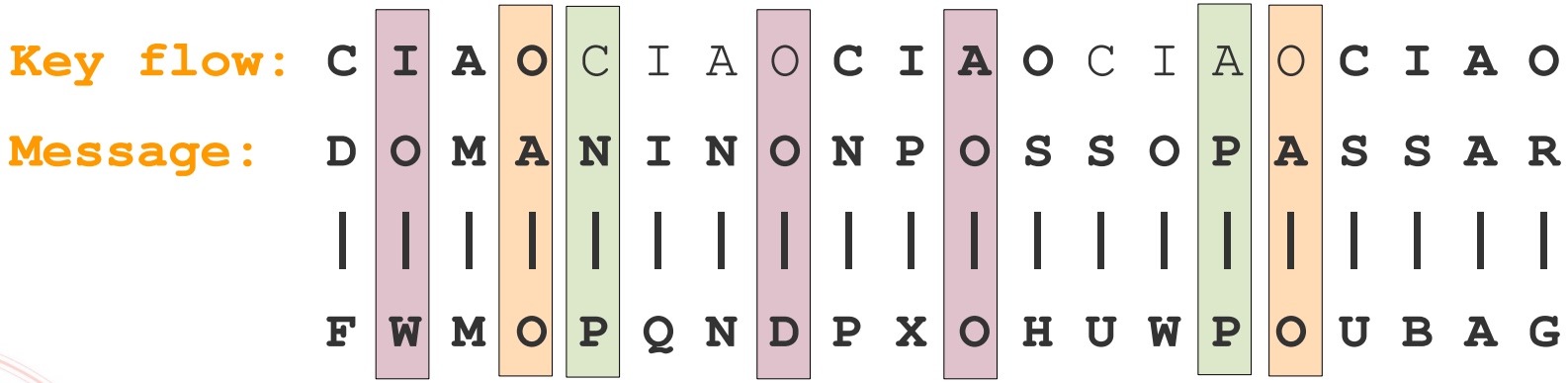
- Leon Battista Alberti (1466)
  - Forma generale e implementazione meccanica
- Bellaso/Vigenère (1553)
  - Forma semplificata usata per 4 secoli (es. la macchina Enigma - WWII)

Sostituzione polialfabetica

Es. si consideri A=0, B=1, … Z=25 e si sommi modulo 26 la chiave al testo

Le frequenze di un carattere in chiaro vengono sparse su più caratteri cifrati

Le frequenze di un carattere cifrato derivano da contributi di diversi caratteri in chiaro



Attaccabile grazie al ripetersi periodico delle sostituzioni

Attaccabile facendo ipotesi sul contenuto del messaggio (*cribs*)
- Trattato sulla crittoanalisi di Charles Babbage (1853)
- Decifrazione rapida di Enigma ad opera di Alan Turing (WWII)

Test crittoanalitici – Kasiski

- Nella polialfabetica la chiave si ripete
- Possono esserci ripetizioni anche di frammenti del testo in chiaro (poligrammi)
- Dove le ripetizioni coincidono → stesso testo cifrato
- Test di Kasiski:
  - ricerca nel cifrato di sequenze identiche
  - annotazione delle distanze
  - fattorizzazione e scelta delle distanze con un fattore comune
  - lunghezza della chiave = MCD
- Esempio:

https://pages.mtu.edu/~shene/NSF-4/Tutorial/VIG/Vig-Kasiski.html

Test crittoanalitici – Indice di coincidenza

- Formalmente, la probabilità che due lettere scelte a caso in un testo siano uguali
- Usato per misurare le variazioni di frequenza delle lettere nel testo cifrato
- Alta variazione è indice di basso "spargimento" → sostituzioni semplici
- Esempio: https://pages.mtu.edu/~shene/NSF-4/Tutorial/VIG/Vig-IOC.html

Es: lunghezza della chiave di un cifrario polialfabetico applicato a un testo naturale

d	IC
1	0.0660
2	0.0520
3	0.0473
4	0.0450
5	0.0436
6	0.0427
7	0.0420
8	0.0415
9	0.0411
10	0.0408
11	0.0405
12	0.0403
13	0.0402
14	0.0400
15	0.0399
...	...
Inf	0.0380

Sost. monoalfabetica – max IC → lingua riconoscibile

distribuzione perfettamente casuale dei caratteri → min. IC … ci si può arrivare?

One-time pad

- Vernam/Mauborgne (1917)
- Polialfabetica con chiave
  - Scelta perfettamente a caso
  - Lunga quanto il messaggio
  - Mai riutilizzata

Ma che fatica!

Testo in chiaro FRA, Testo cifrato: WPE

Tutte equiprobabili

Chiavi possibili AAA … EVT … DYE … **RYE** … FHQ …

Testi in chiaro WPE … SUL … TRA … **FRA** … RIO …

Ipotesi valide: ***tutte*** quelle della lingua considerata

→ Quella giusta è indistinguibile

Sicurezza perfetta!

Have fun

Il “coltellino svizzero” della crittografia

https://gchq.github.io/CyberChef/

Molti esempi di utilizzo avanzato su casi di interesse pratico

https://github.com/mattnotmax/cyberchef-recipes